Journal of Vegetation Science 18: 751-754, 2007
© IAVS; Opulus Press Uppsala.

Improved estimates of incident radiation and heat load using non-parametric regression against topographic variables

McCune, Bruce

Department of Botany & Plant Pathology, Cordley 2082, Oregon State University, Corvallis, OR 97331-2902, USA;
E-mail bruce.mccune@science.oregonstate.edu

Abstract

Question: Can non-parametric multiplicative regression (NPMR) improve estimates of potential direct incident radiation (PDIR) and heat load based on topographic variables, as compared to least-squares multiple regression against trigonometric transforms of the predictors?

Methods: We used a multiplicative kernel smoothing technique to interpolate between tabulated values of PDIR, using a locally linear model and a Gaussian kernel, with slope, aspect, and latitude as predictors. Heat load was calculated as a 45 degree rotation of the PDIR response surface.

Results: This method yielded a fit to a complex response surface with $R^2 > 0.99$ and eliminated the areas of poor fit given by a previously published method based on least squares multiple regression with trigonometric functions of the predictors.

Conclusions: Improved estimates of PDIR and heat load based on topographic variables can be obtained by using non-parametric multiplicative regression (NPMR). The main drawback to the method is that it requires reference to the data tables, since those data are part of the model.

Keywords: Aspect; Azimuth; HyperNiche; Latitude; Light; Non-parametric Multiplicative Regression; NPMR, Slope; Temperature; Topography.

Abbreviations: NPMR = non-parametric multiplicative regression; PDIR = potential direct incident radiation.

Introduction

One way to estimate potential direct incident radiation at various points on the Earth is to simulate solar trajectories, accumulating totals (e.g. Buffo et al. 1972; Iqbal 1983). This method can be applied to calculate representative daily values (e.g. Ejrnæs & Bruun 2000) but it can be cumbersome to apply, requiring numerical integration to obtain yearly totals. A second commonly used method calculates a unitless index of solar radiation based on tan(slope) × cos(aspect-180) (e.g. Stage 1976; Parker 1988). A third method approximates values by regression, based on tables of incident radiation according to slope, aspect, and latitude (e.g. results of Buffo et al. 1972). McCune & Keon (2002) estimated incident radiation and a heat load index by least-squares multiple regression using trigonometric functions of those three predictors. Extension of this approach to the southern hemisphere and errata are in Table 1. While the results of McCune & Keon's equations are conveniently calculated in a spreadsheet, use of the equations has revealed some areas of the regression where the fit is biased. The present paper offers another method for estimating PDIR and heat load using non-parametric regression that provides excellent fit throughout all combinations of the predictors. The only drawback to this approach is that estimates must be made with reference to spreadsheets that contain the original data tables, rather than with an equation.

Table 1. Two errors in the equations for potential annual direct incident radiation and heat load for the northern hemisphere presented by McCune & Keon (2002) and supplemental equations for the southern hemisphere. Equations in McCune & Keon (2002) can be applied to the southern hemisphere by using a different formula for folded aspect. Enter latitudes as positive values. Diagrams on McCune's web site http://oregonstate.edu/~mccuneb/ illustrate the principle. The equations for folding aspect are given below in degrees, so conversion to radians may be needed.

	N hemisphere*	S hemisphere
Folded aspect for incident radiation	180 – | Aspect – 180 |	| Aspect – 180 |
Folded aspect for heat load	| 180 – | Aspect – 225 ||	| 180 – | Aspect – 315 ||

*Errors in the equations in McCune & Keon (2002):
1. Excel equation on p. 605, the last term of the first line should be –1.5*COS(I3), not –1.*COS(I3);
2. In Table 2, p. 605, the third to last value in the "Eq. 1" column should be –0.984, not –0.939.

Non-parametric regression

Given the table of potential direct incident radiation (PDIR) values in relation to slope, aspect, and latitude in Buffo et al. (1972), the problem can be viewed as interpolation for any combination of those predictors. Because the response variable (incident radiation) varies in a nonlinear, interactive way against the three predictors, the underlying function is complex and difficult to specify. Use of a smoothing function to interpolate avoids having to specify the function.

Non-parametric regression estimates a response at a given point in the predictor space by weighting heavily points that are near the target point, and giving little weight to distant points. We used a locally linear model to estimate the response from a weighted least-squares regression, the weights dependent on those distances. A Gaussian function specified smoothly diminishing weights with distance from the target point. Weights from individual variables were combined multiplicatively, automatically accommodating interactions among the predictors – this is non-parametric multiplicative regression (NPMR; McCune 2006). Other smoothers could also be used, for example, generalized additive models (GAMs), but for those, interactions among predictors must be specified explicitly.

NPMR models were fit to Buffo's data with HyperNiche (McCune & Mefford 2004) using a locally linear model, a Gaussian kernel, and a minimum average neighborhood size, $N^* = 3$, where N^* is the average sum of the weights for other data points that bear on the target point. Using $N^*=3$ provides a more flexible response surface than the default suggested by HyperNiche ($N^* = 22$ in this case). Because the data tables in Buffo et al. were generated deterministically, the surface can be fit using a small N^*, with no risk of overfitting.

Sensitivity of the model to each predictor was measured by nudging the values up and down by 5% of the range of individual predictors, and calculating the resulting change in the estimate for that point (McCune 2006). Those differences were averaged over all of the data points. The change in the response can be measured as a fraction of the observed range of the response variable. Scaling the nudgings and responses by the ranges of the variables results in a sensitivity measure that is a ratio, independent of the units of the variables. A sensitivity of 1.0 means that, on average, nudging a predictor results in a change in response of equal magnitude, relative to their ranges. Sensitivity = 0.0 means that nudging a predictor has no detectable effect on the response.

Using the 441 values in Buffo's table of PDIR as the response variable, the final NPMR model with three predictors (Table 2) yielded a cross-validated $xR^2 = 0.994$ (total sum of squares = 37.560; leave-one-out cross-validated residual sum of squares = 0.210), with an average neighborhood size = 4.5.

NPMR estimates of PDIR offer several advantages over the equations in McCune & Keon (2002): (1) the fit is better ($xR^2 = 0.994$ vs. $R^2 = 0.978$; Fig. 1); (2) NPMR removes areas of bias near the peak PDIR levels; (3) NPMR was implemented with angle measurements in degrees, rather than converting to radians; and (4) reflection of Buffo's tables around the north-south axis in the data tables eliminates the need to 'fold' aspects. The disadvantage of the NPMR approach is that it cannot be represented by an equation in a single cell of a spreadsheet.

Both NPMR and the equations in McCune & Keon (2002) are much superior to a radiation index based on tan(slope) × cos(aspect–180) (Fig. 1). Note that 90° slopes were excluded when calculating the latter index to avoid the undefined tan(90°). Even restricting the comparison to a single latitude (40°N) revealed a poor relationship between the radiation index and PDIR (Fig. 1).

Implementation

Estimates for new points are based on a local NPMR model applied to Buffo's data. A local model is fit to each new point, one at a time. The original data are, therefore, an essential part of the model.

Estimates of PDIR for a set of topographic points require the following components:

1. A file with the latitude (0-60°), slope (0-90°), and aspect (0-360° E of N) for a set of points (App. 2).
2. A file with Buffo's data (App. 2, online supplement). The PDIR values for 0-180° aspects have been reflected about the N-S axis, producing values for 180-360°.
3. Specification of the smoothing parameters ('tolerances' in HyperNiche) for each predictor (Table 2).
4. A program to calculate the local estimate (weighted least squares) for each of the points listed in #1.

We implemented the model in HyperNiche for convenience (Table 1), but a similar procedure could be used with other statistical packages offering multidimensional smoothing functions.

Table 2. Key attributes of the predictors in the NPMR model of potential direct incident radiation. Tolerance is the bandwidth used in the multiplicative kernel smoother, given in the units of the predictor; sensitivity is a unitless measure of the importance of a predictor in the model (0 = no response, 1 = 1:1 change in predictor and response, proportionate to ranges).

Predictor	Tolerance	Sensitivity	Range
Latitude, degrees	10.2	0.3748	0 – 60
Slope, degrees	6.3	0.5643	0 – 90
Aspect, degrees	19.8	0.3516	0 - 360

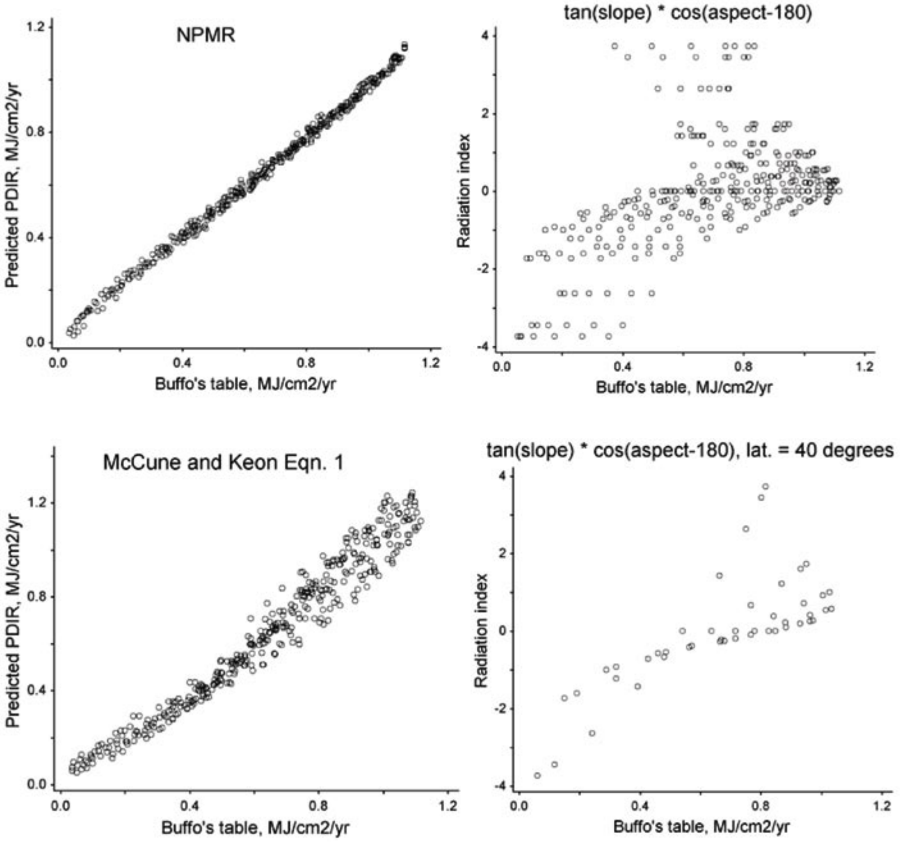

Fig. 1. Predicted versus observed values from regression of potential direct incident radiation (PDIR) against topographic variables. Upper left: Non-parametric multiplicative regression (NPMR). Lower left: Multiple linear regression based on trigonometric functions, from McCune & Keon (2002). Compare these with the commonly used radiation index tan(slope)*cos(aspect-180) (upper right for 0-60°N, lower right for 40°N only).

Table 3. File specifications. Files are provided (online http://oregonstate.edu/~mccuneb/) for estimating potential direct incident radiation (PDIR) and heat load for new sites (Table 1). Two sets of files are available, one for the northern hemisphere and one for the southern hemisphere.

	Northern hemisphere	Southern hemisphere
PDIR	PDIR-N.wk1 833 rows × 1 column	PDIR-S.wk1 833 rows × 1 column
Heat load	HeatLoad-N.wk1 833 rows × 1 column	HeatLoad-S.wk1 833 rows × 1 column
Topographic variables	LatSlopeAspect360.wk1 833 rows × 3 columns	LatSlopeAspect360.wk1 833 rows × 3 columns (same as for N hemisphere)
Model specification file for HyperNiche	BestLLModel.spx (same model file for both PDIR and Heat load)	BestLLModel.spx (same as for N hemisphere)

Row names in these files follow the convention: Rows for aspects from 0-180 degrees east of north are named ObsE1, ObsE2, etc. Rows for aspects from 180-360 degrees east of north are named ObsW1, ObsW2, etc.

In HyperNiche, generate estimates for new points on the landscape with the following steps.

1. Create a spreadsheet with row (point) names, plus three data columns named lat, slope, aspect (all lower case). Note that all angles must be in the domain 0-360 degrees. Convert the data to degrees, if not already in those units. Save as a *.wk1 file.
2. In HyperNiche, open the appropriate PDIR or heat load file as the response matrix.
3. Open the model specification file (*File* | *Open* | *Model file* | BestLLModel.spx).
4. Open the file LatSlopeAspect360 as the predictor matrix.
5. Select *Prediction* | *New sites* and select the file you prepared in step 1. Select other options as desired. The resulting estimates will be written into the result file and/or a spreadsheet, as you choose.

Heat load

Potential direct incident radiation is symmetrical about the north-south axis, but temperatures are not, because a slope with afternoon sun will have higher maximum temperatures than an equivalent slope with morning sun. One can thus convert estimates of PDIR into an approximation of "heat load" by rotating the response surface 45 degrees, such that the index peaks with a SW aspect in the northern hemisphere and is minimized with a NE aspect (Table 1; McCune & Keon 2002).

McCune and Keon (2002) calculated a heat load index by adjusting the equation for folded aspect and otherwise using the same equation as for PDIR. With NPMR, we retain the same smoothing parameters as for PDIR, but rotate the aspects by 45° by shifting the values in the response variable (App. 2). Clearly, this approach has limitations, since the relationship between heat load and aspect is likely to vary by latitude and slope. Given these problems, the crude approximation of heat load by a 45° rotation of PDIR is best restricted to a relatively narrow range of latitudes.

Acknowledgements. Thanks to John Ive and Sean Parks for pointing out problems with the McCune and Keon equations, leading to this alternative method. Thanks also to Sean Parks for implementing a GIS version of the McCune and Keon equations. Patricia Muir and Hans Henrik Bruun helped by critiquing the manuscript.

References

Buffo, J., Fritschen, L.J. & Murphy, J.L. 1972. Direct solar radiation on various slopes from 0 to 60 degrees north latitude. *USDA Forest Service Research Paper* PNW-142. Portland, OR, US.

Ejrnæs, R. & Bruun, H.H. 2000. Gradient analysis of dry grassland vegetation in Denmark. *J. Veg. Sci.* 11: 573-584.

Iqbal, M. 1983. *Introduction to solar radiation.* Academic Press, Toronto, CA.

Klein, S.A. 1977. Calculation of monthly average insolation on tilted surfaces. *Solar Energy* 19: 325-329.

McCune, B. 2006. Non-parametric habitat models with automatic interactions. *J. Veg. Sci.* 17: 819-830.

McCune, B. & Keon, D. 2002. Equations for potential annual direct incident radiation and heat load. *J. Veg. Sci.* 13: 603-606.

McCune, B. & Mefford, M.J. 2004. *HyperNiche. Non-parametric multiplicative habitat modeling*. Version 1.16. MjM Software, Gleneden Beach, OR, US.

Parker, K.C. 1988. Environmental relationships and vegetation associates of columnar cacti in the northern Sonoran Desert. *Vegetatio* 78: 125-140.

Stage, A.R. 1976. An expression for the effect of aspect, slope, and habitat type on tree growth. *For. Sci.* 4: 457-460.

Received 1 January 2007;
Accepted 22 February 2007;
Co-ordinating Editor: R. Ejrnæs.

Tables 1 and 3 also available on line as Electronic Appendices 1 and 2 at www.opuluspress.se